SORT 29 (2) July-December 2005, 217-234

On sequential and fixed designs for estimation with comparisons and applications

Mekki Terbeche[1], Broderick O. Oluyede[2], Ahmed Barbour[3]

[1]University of Oran, [2]Georgia Southern University, [3]U.A.E. University

Abstract

A fully sequential approach to the estimation of the difference of two population means for distributions belonging to the exponential family of distributions is adopted and compared with the best fixed design. Results on the lower bound for the Bayes risk due to estimation and expected cost are presented and shown to be of first order efficiency. Applications involving the Poisson and exponential distributions with gamma priors as well as the Bernoulli distribution with beta priors are given. Finally, some numerical results are presented.

MSC: 62L12

Keywords: Estimation, Loss function, Prior distribution, Sequential design.

1 Introduction

The family of exponential type distributions play an important role in a wide variety of areas in probability and statistics. For example, the gamma distribution which belong to the family of exponential distributions is used to model lifetimes of various practical situations including but not limited to lengths of time between catastrophic events (floods, earthquakes and so on), lengths of time between emergency arrivals at a hospital and distance traveled by a wildlife ecologist between sighting of an endangered species. The exponential distribution which is a special case of the gamma distribution have

Address for correspondence: Mekki Terbeche. Department of Mathematics, University of Oran, Es-Senia, Algeria. Broderick O. Oluyede. Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA 30460. *Boluyede@GeorgiaSouthern.edu. Ahmed Barbour. College of Information Technology, U.A.E. University, Al Ain, United Arab Emirates.

Received: June 2001
Accepted: June 2005

been used to describe the amount of time between occurrences of random events such as those described above. Further examples of the exponential type distributions include the Poisson and binomial distributions. The Poisson distribution provides a realistic model for many random phenomena such as number of fatal traffic accident per week at a busy intersection, the number of radioactive particle emissions per unit time, the number of telephone calls per hour arriving at a switchboard to mention a few. In this paper, we consider the problem of designing an experiment to estimate the difference between two population means for distributions belonging to the exponential family plus expected cost of drawing samples from either groups using a Bayesian approach. We explore and compare the Bayes risk due to estimation plus the expected cost of sampling. Numerical results on the relative efficiency are also presented.

This paper is organized as follows. Section 2 contains some preliminaries and basic results for the class of exponential type distributions. In Section 3, the problem is presented and the mathematical results on the fully sequential and best fixed designs derived. Some bounds are presented. In Section 4, we present applications and numerical results on the comparisons of the Bayes risk for the procedures described in Section 3. Applications are presented for the comparisons of two Poisson means, comparisons of two exponential means and the comparison of two Bernoulli means. Applications on the comparisons of the normal means with known variances as well as the comparisons of two normal variances with known means will be treated in the future. This paper concludes with a summary and discussion.

2 Preliminaries and basic results

In this section, we consider the family of exponential-type probability distributions on the real line, given by the family of densities $\mathcal{G}$ with respect to the Lebesgue measure. A natural form of an exponential family is as follows:

$$f(x, \theta) = \exp\{\theta T(x) + S(x) - \psi(\theta)\}, \tag{1}$$

where $f \in \mathcal{G}$. In this setting $E(T(X)) = \psi^{'}(\theta)$ and $\text{var}(T(X)) = \psi^{''}(\theta)$. See Lehmann [3]. Consider two independent random variables X and Y with densities given by $f(x, \theta) = \exp\{\theta T(x) + S(x) - \psi(\theta)\}$, and $g(y, \omega) = \exp\{\omega T(y) + U(y) - \phi(\omega)\}$ respectively. Our objective is to estimate $\lambda = E_\theta[T(X)] - E_\omega[T(X)] = \psi^{'}(\theta) - \phi^{'}(\omega)$ with square error loss.

Definition 1 *The Bayes risk of an estimate $\hat{\lambda}$ with respect to the prior distribution $\pi(\theta)$ is*

$$r(\theta, \hat{\lambda}) = E[R(\theta, \hat{\lambda})], \tag{2}$$

where $R(\theta, \hat{\lambda}) = E[L(\theta, \hat{\lambda})]$ and $L(\theta, \hat{\lambda})$ is the loss function.

The adopted approach in this paper is Bayesian and it is assumed that the prior distributions of θ and ω are the conjugate priors given by: $\pi_1(\theta) \propto \exp[t(\theta\mu_1 - \phi(\theta))]$ and $\pi_2(\omega) \propto \exp[s(\omega\mu_2 - \phi(\omega))]$, where $\mu_1 = E_{\pi_1}[\phi^{'}(\theta)]$ and $\mu_2 = E_{\pi_2}[\psi^{'}(\omega)]$ are prior estimations of $E_\theta[T(X)]$ and $E_\omega[T(X)]$ respectively, if these densities and their derivatives decay to zero in the tails, (See West, 1985, 1986), and $t > 0$ and $s > 0$ are positive real numbers that can be interpreted as prior sample sizes.

If $X_1, X_2, \ldots, X_m$ is a random sample of X and $Y_1, Y_2, \ldots, Y_n$ is a random sample of Y, the Bayes estimator of λ is given by

$$\begin{aligned} \hat{\lambda}(x_1, \ldots, x_m, y_1, \ldots, y_n) &= E[\lambda | x_1, \ldots, x_m, y_1, \ldots, y_n] \\ &= E[\psi^{'}(\theta)|x_1, \ldots, x_m] - E[\psi^{'}(\omega)|y_1, \ldots, y_n], \end{aligned} \tag{3}$$

where

$$E[\psi^{'}(\theta)|x_1, \ldots, x_m] = \frac{m\overline{T}_m^X + t\mu_1}{m + t}$$

with $\overline{T}_m^X = \frac{T(x_1)+\cdots+T(x_m)}{m}$ and

$$E[\psi^{'})(\omega)|y_1, \ldots, y_n] = \frac{m\overline{T}_n^Y + s\mu_2}{n + s}$$

with $\overline{T}_n^Y = \frac{T(y_1)+\cdots+T(y_n)}{n}$.

If $\mathbf{X}=(X_1, \ldots, X_m)$ and $\mathbf{Y}=(Y_1, \ldots, Y_n)$, $\mathbf{x}=(x_1, \ldots, x_m)$ and $\mathbf{y}=(y_1, \ldots, y_n)$, the Bayes risk is given by

$$\begin{aligned} r(\pi_1, \pi_2) &= r(\hat{\lambda}(\mathbf{x}, \mathbf{y})) \\ &= E_{(\mathbf{X},\mathbf{Y})}\left[E_{\lambda|(\mathbf{X},\mathbf{Y})}\left[(\lambda - \hat{\lambda}(\mathbf{x}, \mathbf{y}))^{\mathbf{2}}\right]\right] \\ &= E_{(\mathbf{X},\mathbf{Y})}[\mathrm{var}(\lambda|(\mathbf{X}, \mathbf{Y}))] \\ &= E_{(\mathbf{X},\mathbf{Y})}[\mathrm{var}(\psi^{'}(\theta)|\mathbf{X}) + \mathrm{var}(\phi^{'}(\omega)|\mathbf{Y})] \\ &= E_{\mathbf{X}}\left[E_{\theta|\mathbf{X}}\left[\frac{\psi^{''}(\theta)}{m + t}\right]\right] + E_{\mathbf{Y}}\left[E_{\omega|\mathbf{Y}}\left[\frac{\phi^{''}(\omega)}{n + s}\right]\right]. \end{aligned} \tag{4}$$

3 Sequential and best fixed designs

3.1 The problem

In order to set up the problem we adopt the notation given in Berger (1985, Chapter 7). The loss function is given by

$$L(\lambda, a, m, n) = (\lambda - a)^2 + c_1 m + c_2 n, \tag{5}$$

and the decision rule are sequential decision procedures $\Delta_S = (\tau, \delta)$ where τ is called the stopping rule of the procedure and consist of functions $\tau_{m,n}(x_1, \ldots, x_m, y_1, \ldots, y_n)$ that specify the probability of stopping sampling and making a decision after observing $(x_1, \ldots, x_m, y_1, \ldots, y_n)$; δ is the decision rule of the design Δ_S and consists of a series of decision functions $\delta_{m,n}(x_1, \ldots, x_m, y_1, \ldots, y_n)$ that specify the estimated value of λ when the sampling has stopped after observing $(x_1, \ldots, x_m, y_1, \ldots, y_n)$.

For the stopping rule τ, the Bayes risk is given by:

$$\begin{aligned} r(\tau, \pi_1, \pi_2) &= E_{(\mathbf{X},\mathbf{Y},\tau)}\left[\frac{U_m}{m+t} + \frac{V_n}{n+s} + c_1 m + c_2 n\right] \\ &= E_{(\mathbf{X},\tau_\mathbf{m})}\left[\frac{U_m}{m+t} + c_1 m\right] + E_{(\mathbf{Y},\tau_\mathbf{n})}\left[\frac{V_n}{n+s} + c_2 n\right], \end{aligned} \tag{6}$$

where τ_m and τ_n are the marginal stopping rules of τ, and $U_m = E_{(\mathbf{X},\tau_\mathbf{m})}[\psi''(\theta)]$, $V_n = E_{(\mathbf{Y},\tau_\mathbf{n})}[\phi''(\omega)]$, t and s are fixed and depend on the posteriors, m and n are unknown.

The fixed designs are particular cases of Δ_S where the stopping rules τ_m and τ_n are equal to one if $m = m^F$ and $n = n^F$ and zero otherwise and their optimal values $m_{opt}(\pi)$ and $n_{opt}(\pi)$ are given in this section.

3.2 Mathematical results

In this subsection, the mathematical results are presented. We compare the best fixed design with the sequential optimal random design. Let c_1 and c_2 be the cost of sampling per observation from populations 1 and 2 respectively. The Bayes risk due to estimation plus expected sampling cost is given by equation (6). The objective or goal is to minimize $r(\tau, \pi_1, \pi_2)$.

In the sequential allocation, for a fixed total number of observations the problem is to allocate the number of observations to be taken from each population to achieve or nearly achieve some optimality condition such as minimizing the Bayes risk when the allocation is done sequentially. That is, at each stage the decision to observe X or Y may depend on available information from all previous stages.

Note that at stage t:

$a)$ If $U_m^{1/2} \geq c_1^{1/2}(m+t)$ take another observation of X; otherwise stop observing X.
$b)$ If $V_n^{1/2} \geq c_2^{1/2}(n+s)$ take another observation of Y; otherwise stop observing X.

The rule takes an additional observation of X (respectively Y) if $m_{opt}(\theta|x_1,\ldots,x_m) \geq 1$ (respectively $n_{opt}(\theta|y_1,\ldots,y_n) \geq 1$, where $m_{opt}(\pi) = \left(E_\pi[\psi^{''}(\theta)]/c_1\right)^{1/2} - t$ (respectively $n_{opt}(\pi) = \left(E_\pi[\phi^{''}(\omega)]/c_2\right)^{1/2} - s$) are the sample sizes of the fixed design when the distribution of θ (respectively ω) is π. The sequential design achieves the lower bound. That is,

$$\liminf_{c_1,c_2\to 0}\left(\frac{r(\Delta)}{(c_1+c_2)^{1/2}}\right) = 2E[(\gamma_1\psi^{''}(\theta))^{1/2} + (\gamma_2\phi^{''}(\omega))^{1/2}]. \tag{7}$$

To see this, and for simplicity of the computations, we take the exponential family with probability distribution of the form $f_\theta(x) = \exp[\theta x - \psi(\theta)]$, $x \in \mathbb{R}$, $\theta \in \Omega$. Clearly, $E_\theta(X) = \psi^{'}(\theta)$ and $\mathrm{var}_\theta(X) = \psi^{''}(\theta)$, after differentiating the identity $\int e^{\theta x - \psi(\theta)}dx = 1$, once and twice with respect to θ and simplifying each expression respectively. Similarly, $E_\omega(Y) = \phi^{'}(\omega)$ and $\mathrm{var}_\omega(Y) = \phi^{''}(\omega)$. Following Diaconis and Ylvisaker [2], the form of the conjugate for exponential families, for $t > 0$ and $s > 0$ are

$$\pi(\theta) = \frac{e^{t[\mu\theta - \psi(\theta)]}}{\int e^{t[\mu\theta-\psi(\theta)]}d\theta}, \tag{8}$$

and

$$\gamma(\omega) = \frac{e^{s[\mu\omega - \phi(\omega)]}}{\int e^{s[\mu\omega-\phi(\omega)]}d\omega}, \tag{9}$$

respectively. We assume that θ and ω are independent random variables with conjugate prior distributions given above. If $(X_1, X_2, \ldots, X_m)$ is a random sample of X and $(Y_1, Y_2, \ldots, Y_n)$ is a random sample of Y, then

$$f_\theta(X_1,\ldots,X_m) = \exp[m(\theta\overline{X} - \psi(\theta))], \tag{10}$$

where $\overline{X} = (X_1,\ldots,X_m)/m$ and

$$g_\omega(Y_1,\ldots,Y_n) = \exp[n(\omega\overline{Y} - \phi(\omega))], \tag{11}$$

where $\overline{Y} = (Y_1,\ldots,Y_n)/n$.

The posterior distribution of θ when m observations $(X_1, X_2, \ldots, X_m)$ are sampled from population 1 is

$$\begin{aligned}\pi(\theta|X_1, X_2, \ldots, X_m) &= f_\theta(X_1, \ldots, X_m) / \int f_\theta(X_1, \ldots, X_m)\pi(\theta)d\theta \\ &= \frac{e^{m(\theta\overline{X}-\psi(\theta))}\frac{e^{t[\mu\theta-\psi(\theta)]}}{\int e^{t[\mu\theta-\psi(\theta)]}d\theta}}{\int e^{m(\theta\overline{X}-\psi(\theta))}\frac{e^{t[\mu\theta-\psi(\theta)]}}{\int e^{t[\mu\theta-\psi(\theta)]}d\theta}d\theta} \\ &= \frac{e^{\theta[m\overline{X}+t\mu]-(m+t)\psi(\theta)}}{\int e^{\theta[m\overline{X}+t\mu]-(m+t)\psi(\theta)}d\theta}. \qquad (12)\end{aligned}$$

Set $t_1 = m + t$ and $\mu_1 = \frac{t}{t_1}\mu + \frac{m}{t_1}\overline{X}$. Then

$$\pi(\theta|X_1, \ldots, X_m) = \frac{e^{t_1[\mu_1\theta-\psi(\theta)]}}{\int e^{t_1[\mu_1\theta-\psi(\theta)]}d\theta}. \qquad (13)$$

Similarly,

$$\gamma(\omega|Y_1, \ldots, Y_n) = \frac{e^{s_1[\mu_2\omega-\phi(\omega)]}}{\int e^{s_1[\nu_1\omega-\phi(\omega)]}d\omega}, \qquad (14)$$

where $s_1 = n + s$ and $\mu_2 = \frac{s}{s_1}\nu + \frac{n}{s_1}\overline{Y}$.

Next we show that the posterior mean and variance of $\psi^{'}(\theta)$ given $x_1, \ldots, x_m$ are $E[\psi^{'}(\theta)|x_1, \ldots, x_m] = \mu_1$ and $\text{var}[\psi^{'}(\theta)|x_1, \ldots, x_m] = E_{\theta|\mathbf{X}}[\frac{\psi^{''}(\theta)}{m+t}]$ respectively. First we state a useful lemma. For a proof of the lemma see Hajek and Sidak (1967).

Lemma 1 *If f is an absolutely continuous integrable and real valued function for which $\int |f(\omega)|d\omega < \infty$, then* $\lim_{\omega\to-\infty} f(\omega) = 0$ *and* $\lim_{\omega\to+\infty} f(\omega) = 0$.

The posterior mean of $\psi^{'}(\theta)$ given $x_1, \ldots, x_m$ is

$$\begin{aligned}E[\psi^{'}(\theta)|x_1, \ldots, x_m] &= \int \psi^{'}(\theta)\pi(\theta|X_1, \ldots, X_m)d\theta \\ &= \int \psi^{'}(\theta)\frac{e^{t[\mu\theta-\psi(\theta)]}}{\int e^{t[\mu\theta-\psi(\theta)]}d\theta}d\theta \\ &= \frac{1}{\int e^{t[\mu\theta-\psi(\theta)]}d\theta}\int \psi^{'}(\theta)e^{t[\mu\theta-\psi(\theta)]}d\theta \\ &= -\frac{1}{t_1\int e^{t[\mu\theta-\psi(\theta)]}d\theta}\int [t_1(\mu_1 - \psi^{'}(\theta))e^{t[\mu\theta-\psi(\theta)]} - t_1\mu_1 e^{t[\mu\theta-\psi(\theta)]}]d\theta\end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{t_1 \int e^{t[\mu\theta-\psi(\theta)]}d\theta} \int \frac{d}{d\theta}[e^{t[\mu\theta-\psi(\theta)]}]d\theta + \mu_1 \int \frac{e^{t[\mu\theta-\psi(\theta)]}}{\int e^{t[\mu\theta-\psi(\theta)]}d\theta}d\theta \\
&= -\frac{1}{t_1 \int e^{t[\mu\theta-\psi(\theta)]}d\theta} \int \frac{d}{d\theta}[e^{t[\mu\theta-\psi(\theta)]}]d\theta + \mu_1 \\
&= \mu_1 - \frac{1}{t_1 \int e^{t[\mu\theta-\psi(\theta)]}d\theta}[\lim_{\theta\to+\infty} e^{t[\mu\theta-\psi(\theta)]} - \lim_{\theta\to+\infty} e^{t[\mu\theta-\psi(\theta)]}] \\
&= \mu_1, \qquad (15)
\end{aligned}
$$

since the two limits vanish by virtue of the lemma given above.

The posterior variance of $\psi^{'}(\theta)$ given $x_1, \ldots, x_m$ is

$$
\begin{aligned}
\mathrm{var}[\psi^{'}(\theta)|x_1,\ldots,x_m] &= \int [\psi^{'}(\theta) - E(\psi^{'}(\theta))]^2 \pi(\theta|X_1,\ldots,X_m)d\theta \\
&= \frac{1}{\int e^{t[\mu\theta-\psi(\theta)]}d\theta} \int [\psi^{'}(\theta) - E(\psi^{'}(\theta))]^2 e^{t[\mu\theta-\psi(\theta)]}d\theta. \qquad (16)
\end{aligned}
$$

Let $\alpha(\theta) = t_1[\mu\theta - \psi(\theta)]$, then

$$
[E(\psi^{'}(\theta)) - \psi^{'}(\theta)]^2 = [\frac{1}{t_1}\frac{d\alpha(\theta)}{d\theta}]^2, \qquad (17)
$$

so that

$$
\begin{aligned}
\mathrm{var}[\psi^{'}(\theta)|x_1,\ldots,x_m] &= \frac{1}{t_1^2 \int e^{t[\mu\theta-\psi(\theta)]}d\theta} \int [\frac{d\alpha(\theta)}{d\theta}]^2 e^{\alpha(\theta)}d\theta \\
&= \frac{1}{t_1^2 \int e^{t[\mu\theta-\psi(\theta)]}d\theta}[[\alpha^{'}(\theta)e^{\alpha(\theta)}]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \alpha^{''}(\theta)e^{\alpha(\theta)}d\theta] \\
&= \frac{1}{t_1^2 \int e^{t[\mu\theta-\psi(\theta)]}d\theta}[\lim_{\theta\to+\infty} \alpha^{'}(\theta)e^{\alpha(\theta)}] - [\lim_{\theta\to-\infty} \alpha^{'}(\theta)e^{\alpha(\theta)}] \\
&- \int_{-\infty}^{\infty} \alpha^{''}(\theta)e^{\alpha(\theta)}d\theta \\
&= -\frac{1}{t_1^2 \int e^{t[\mu\theta-\psi(\theta)]}d\theta} \int_{-\infty}^{\infty} \alpha^{''}(\theta)e^{\alpha(\theta)}d\theta \\
&= \frac{t_1}{t_1^2 \int e^{t[\mu\theta-\psi(\theta)]}d\theta} \int \psi^{''}(\theta)e^{t[\mu\theta-\psi(\theta)]}d\theta
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{t_1}\int \frac{\psi^{''}(\theta)e^{t[\mu\theta-\psi(\theta)]}}{\int e^{t[\mu\theta-\psi(\theta)]}d\theta}d\theta \\
&= \frac{1}{t_1}E[\psi^{''}(\theta)|x_1,\ldots,x_m] \\
&= \frac{E[\psi^{''}(\theta)|x_1,\ldots,x_m]}{m+t} \\
&= E[\frac{\psi^{''}(\theta)}{m+t}|x_1,\ldots,x_m] \\
&= E_{\theta|\mathbf{X}}[\frac{\psi^{''}(\theta)}{m+t}],
\end{aligned}
\tag{18}
$$

which also gives the proof of equation 4. □

In the best fixed design or policy the risk function $r(\Delta)$ is minimized as a function of fixed sample sizes m and n. This policy is asymptotically the best among the non-sequential or non-random policies. The best fixed design is determined by $m_{opt}(\pi) = (E[\psi^{''}(\theta)]/c_1)^{1/2} - t$ and $n_{opt}(\pi) = (E[\phi^{''}(\omega)]/c_2)^{1/2} - s$, and achieves the lower bound under suitable conditions.

Theorem 1 *Let c_1 and c_2 be such that $\frac{c_1}{c_1+c_2} \to \gamma_1$, as c_1, $c_2 \to 0$, $0 < \gamma_1 < 1$ and $\gamma_2 = 1 - \gamma_1$. Then for any random design Δ,*

$$\liminf_{c_1,c_2\to 0}\left(\frac{r(\Delta)}{(c_1+c_2)^{1/2}}\right) \geq 2E[(\gamma_1\psi^{''}(\theta))^{1/2} + (\gamma_2\phi^{''}(\omega))^{1/2}].$$

Proof. Observe that

$$r(\Delta) \geq 2E\left[(c_1U_m)^{1/2} + (c_2V_n)^{1/2}\right] - tc_1 - sc_2. \tag{19}$$

for any procedure Δ.

Now,

$$
\begin{aligned}
\left(\frac{r(\Delta)}{(c_1+c_2)^{1/2}}\right) &\geq 2E\left[\frac{(c_1U_m)^{1/2}}{(c_1+c_2)^{1/2}} + \frac{(c_2V_n)^{1/2}}{(c_1+c_2)^{1/2}}\right] \\
&- tc_1^{1/2}\left(\frac{c_1}{c_1+c_2}\right)^{1/2} - sc_2^{1/2}\left(\frac{c_2}{c_1+c_2}\right)^{1/2}.
\end{aligned}
\tag{20}
$$

Consequently,

$$\liminf_{c_1,c_2\to 0}\left(\frac{r(\Delta)}{(c_1+c_2)^{1/2}}\right) \geq 2E[(\gamma_1\psi''(\theta))^{1/2} + (\gamma_2\phi''(\omega))^{1/2}].$$

The last inequality follows from the application of Fatou's lemma.

Note that for any fixed design $\Delta_{\mathcal{F}}$,

$$\begin{aligned} r(\Delta_{\mathcal{F}}) &= 2[(c_1E_\theta[\psi''(\theta)])^{1/2} + (c_2E_\omega[\psi''(\omega)])^{1/2}] \\ &+ (m+t)^{-1}[E_\theta(\psi''(\theta))^{1/2} - (m+t)c_1^{1/2}]^2 \\ &+ (n+s)^{-1}[E_\omega(\phi''(\omega))^{1/2} - (n+s)c_2^{1/2}]^2 - (tc_1+sc_2). \end{aligned} \quad (21)$$

If $m = \left(E_\pi[\psi''(\theta)]/c_1\right)^{1/2} - t$ and $n = \left(E_\pi[\phi''(\omega)]/c_2\right)^{1/2} - s$, then

$$r(\Delta_{\mathcal{F}}) = 2E[(c_1E_\theta[\psi''(\theta)])^{1/2} + (c_2E_\omega[\phi''(\omega)])^{1/2}] - (tc_1+sc_2). \quad (22)$$

Moreover, if c_1 and c_2 are such that $\frac{c_1}{c_1+c_2} \to \gamma_1$, as $c_1, c_2 \to 0$, $0 < \gamma_1 < 1$ and $\gamma_2 = 1-\gamma_1$, then

$$\liminf_{c_1,c_2\to 0}\left(\frac{r(\Delta_{\mathcal{F}})}{(c_1+c_2)^{1/2}}\right) = 2[(\gamma_1E\theta[\psi''(\theta)])^{1/2} + (\gamma_2E_\omega[\phi''(\omega)])^{1/2}].$$

□

Theorem 2 *Let $\Delta_{\mathcal{S}}$ and $\Delta_{\mathcal{F}}$ denote the first order sequential and fixed designs respectively. Then*

$$0 \leq \liminf_{c_1,c_2\to 0}\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} \leq 1. \quad (23)$$

Proof. Note that

$$\liminf_{c_1,c_2\to 0}\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = \frac{(\gamma_1)^{1/2}E(\psi''(\theta))^{1/2} + (\gamma_2)^{1/2}E(\varphi''(\omega))^{1/2}}{(\gamma_1E\psi''(\theta))^{1/2} + (\gamma_2E\varphi''(\omega))^{1/2}}. \quad (24)$$

Applying Jensen's inequality, we have

$$0 \leq \liminf_{c_1,c_2\to 0}\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} \leq 1. \quad (25)$$

□

Theorem 3 *If $c_1 = c_2$ then*

$$\liminf_{c_1 \to 0} \frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = \frac{E(\psi''(\theta))^{1/2} + E(\varphi''(\omega))^{1/2}}{(E\psi''(\theta))^{1/2} + (E\varphi''(\omega))^{1/2}}. \tag{26}$$

Proof. We have

$$\liminf_{c_1, c_2 \to 0} \frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = \frac{(\gamma_1)^{1/2} E(\psi''(\theta))^{1/2} + (\gamma_2)^{1/2} E(\varphi''(\omega))^{1/2}}{(\gamma_1 E\psi''(\theta))^{1/2} + (\gamma_2 E\varphi''(\omega))^{1/2}}. \tag{27}$$

If $c_1 = c_2$, then $\gamma_1 = \gamma_2$ and the result follows. □

Corollary 1 *If $c_1 = c_2$, $\psi''(\theta) = \varphi''(\omega)$, and $\pi_1 = \pi_2$, then*

$$\liminf_{c_1 \to 0} \frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = \frac{E(\psi''(\theta))^{1/2}}{(E\phi''(\omega))^{1/2}}. \tag{28}$$

□

The asymptotic results in section 3 are derived in the following sense. Sampling sizes tending to infinity are achieved by taking the costs of sampling (c_1, c_2) tending to zero, since c_1 and c_2 may differ from population to population. Simultaneous control over c_1 and c_2 is maintained by assuming that $\frac{c_1}{c_1+c_2} \to \gamma_1$, $\frac{c_2}{c_1+c_2} \to \gamma_2$, so that c_1 and c_2 tend to zero at the same rate.

Theorem 2 states that the lower bound for the sequential design is smaller than the lower bound for the best fixed design. This makes sense due to the fact that we use all previous information about the population for the sequential design as well as the information on the priors for the best fixed design.

4 Application

In this section, we present applications of the results in Sections 2 and 3. Specifically, applications involving the Poisson and exponential distributions with gamma priors as well as the Bernoulli distribution with beta priors are given. Some numerical results on the relative efficiency of the estimation problem concerning the Poisson and exponential distributions with gamma priors are also presented.

4.1 Comparison of two Poisson means

Let the distribution of the random variables X and Y be given by $f(x, \theta)$ and $g(y, \omega)$ respectively, where

$$f(x,\theta) = \frac{\theta^x e^{-\theta}}{x!}, \tag{29}$$

$x = 0, 1, 2, \ldots, \theta > 0$ and

$$g(y,\omega) = \frac{\omega^y e^{-\omega}}{y!}, \tag{30}$$

$y = 0, 1, 2, \ldots, \omega > 0$. We assume that θ and ω are independent and distributed as $Gamma(a, p)$, $a > 0$, $p > 0$ and $Gamma(c, q)$, $c > 0$, $q > 0$. It follows therefore from Theorem 1 that

$$\liminf_{c_1,c_2 \to 0} \frac{R(\Delta_{\mathcal{S}})}{R(\Delta_{\mathcal{F}})} = \frac{\left(\frac{\Gamma(a+1/2)}{p^{1/2}\Gamma(a)} + \frac{\Gamma(c+1/2)}{q^{1/2}\Gamma(c)}\right)}{((a/p)^{1/2} + (c/q)^{1/2})}, \tag{31}$$

$a > 0, c > 0, p > 0, q > 0.$ □

Note that (a) If $a/p = c/q = d > 0$, then

$$\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = (1/2)\left(\frac{\Gamma(a+1/2)}{a^{1/2}\Gamma(a)} + \frac{\Gamma(c+1/2)}{c^{1/2}\Gamma(c)}\right). \tag{32}$$

(b) If $a \to 0$ and $c \to \infty$, then

$$\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} \to \frac{1}{2}. \tag{33}$$

(c) If $a = c$ and $p = q$, then

$$\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = \frac{\Gamma(a+1/2)}{a^{1/2}\Gamma(a)}. \tag{34}$$

(d) If $a = c = p = q = 1/2$, then

$$\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = 0.7979. \tag{35}$$

□

4.2 Comparison of two Bernoulli means

Let the distribution of the random variables X and Y be given by $f(x,\theta)$ and $g(y,\omega)$ respectively, where

$$f(x,\theta) = \theta^x(1-\theta)^{1-x}, \tag{36}$$

$x = 0, 1,\ 0 < \theta < 1$ and

$$g(y,\omega) = \omega^y(1-\omega)^{1-y}, \tag{37}$$

$y = 0, 1,\ 0 < \omega < 1$. We assume that θ and ω are independent and distributed as $Beta(a,b)$, $a > 0$, $b > 0$ and $Beta(c,d)$, $c > 0$, $d > 0$. It follows therefore from Theorem 3 that

$$\liminf_{c_1,c_2\to 0} \frac{R(\Delta_{\mathcal{S}})}{R(\Delta_{\mathcal{F}})} = \frac{E(\theta(1-\theta))^{1/2} + E(\omega(1-\omega))^{1/2}}{(E[\theta(1-\theta)])^{1/2} + (E[\omega(1-\omega)])^{1/2}}, \tag{38}$$

where

$$E(\theta(1-\theta))^{1/2} = \frac{\Gamma(a+1/2)\Gamma(b+1/2)}{(a+b)\Gamma(a)\Gamma(b)}, \tag{39}$$

for $a > 0$, $b > 0$, and

$$E(\omega(1-\omega))^{1/2} = \frac{\Gamma(c+1/2)\Gamma(d+1/2)}{(c+d)\Gamma(c)\Gamma(d)}, \tag{40}$$

for $c > 0$, $d > 0$. Similarly,

$$E[\theta(1-\theta)] = ab/(a+b+1)(a+b), \tag{41}$$

for $a > 0$, $b > 0$, and

$$E[\omega(1-\omega)] = cd/(c+d+1)(c+d), \tag{42}$$

for $c > 0$, $d > 0$. For the beta distribution, that is, $\theta \sim Beta(a,b)$, it is well known that $E(\theta) = \frac{a}{a+b}$, and $\mathrm{var}(\theta) = \frac{ab}{(a+b+1)(a+b)^2}$. Similarly, if $\omega \sim Beta(c,d)$, then $E(\omega) = \frac{c}{c+d}$, and $\mathrm{var}(\omega) = \frac{cd}{(c+d+1)(c+d)^2}$.

The ratio of the sequential to the best fixed design is

$$\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = \frac{\dfrac{\Gamma(a+1/2)\Gamma(b+1/2)}{(a+b)\Gamma(a)\Gamma(b)} + \dfrac{\Gamma(c+1/2)\Gamma(d+1/2)}{(c+d)\Gamma(c)\Gamma(d)}}{(ab/(a+b+1)(a+b))^{1/2} + (cd/(c+d+1)(c+d))^{1/2}}, \tag{43}$$

$a > 0$, $b > 0$, $c > 0$, $d > 0$.

Note that (a) If $a = c$ and $b = d$, then

$$\liminf_{c_1, c_2 \to 0} \frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = \frac{\Gamma(a + 1/2)\Gamma(b + 1/2)}{a^{1/2}\Gamma(a)b^{1/2}\Gamma(b)}. \tag{44}$$

(b) For any fixed b

$$\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} \to \frac{\Gamma(b + 1/2)}{b^{1/2}\Gamma(b)}, \tag{45}$$

as $a \to \infty$, and as $a, b \to \infty$

$$\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} \to 1, \tag{46}$$

(c) If $a = b = c = d$, then

$$\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = \left(\frac{\Gamma(a + 1/2)}{a^{1/2}\Gamma(a)}\right)^2 \left(\frac{2a + 1}{2a}\right)^{1/2}. \tag{47}$$

(d) If $a = b = c = d = 1$, then

$$\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = [\Gamma(3/2)]^2(3/2)^{1/2} = 0.9619. \tag{48}$$

(e) If $a, b \to 0$ then

$$\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} \to 0. \tag{49}$$

□

4.3 Comparison of two exponential means

We next consider the estimation of the difference of the means of two exponential populations with gamma priors. Let the distribution of X and Y be given by

$$f(x, \theta) = \theta e^{-\theta x}, \tag{50}$$

for $x > 0, \theta > 0$ and

$$g(y, \omega) = \omega e^{-\omega y}, \tag{51}$$

for $y > 0$, $\omega > 0$ respectively. We assume the prior distributions are $Gamma(a, p)$, $a > 2$,

$p > 0$ and $Gamma(c, q)$, $c > 2$, $q > 0$ respectively. The ratio of the sequential to the best fixed design is given by

$$\liminf_{c_1,c_2\to 0} \frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = \frac{E(\theta^{-1}) + E(\omega^{-1})}{(E(\theta^{-2}))^{1/2} + (E(\omega^{-2}))^{1/2}}. \tag{52}$$

Therefore the ratio of the sequential to the best fixed design becomes

$$\liminf_{c_1,c_2\to 0} \frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = \frac{\dfrac{p}{a-1} + \dfrac{q}{c-1}}{\dfrac{p}{((a-1)(a-2))^{1/2}} + \dfrac{q}{((c-1)(c-2))^{1/2}}}, \tag{53}$$

$a > 2$, $p > 0$, $c > 2$, $q > 0$.

If $a = p$ and $q = c$ then the ratio becomes

$$\liminf_{c_1,c_2\to 0} \frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = \frac{\dfrac{a}{a-1} + \dfrac{c}{c-1}}{\dfrac{a}{((a-1)(a-2))^{1/2}} + \dfrac{c}{((c-1)(c-2))^{1/2}}}, \tag{54}$$

$a > 2$, and $c > 2$.

4.4 Numerical comparisons

In this section we examine the ratio of the sequential to the best fixed designs for the estimation problem. We consider the case of balanced and unbalanced designs. This numerical study is conducted for the case of exponential distribution means with gamma priors and Poisson distribution means with gamma priors. For the balanced designs, $E(\theta) = E(\omega)$ and $\text{var}(\theta) = \text{var}(\omega)$, that is $a = c$ and $p = q$. Note that for the Poisson means with gamma priors with $a/p = c/q = k$, where $k > 0$ is fixed, the ratio $\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})}$ is given by

$$\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})} = \frac{\Gamma(a+0.5)}{2a^{1/2}\Gamma(a)} + \frac{\Gamma(c+0.5)}{2c^{1/2}\Gamma(c)}. \tag{55}$$

Table 1 gives the $\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})}$ for the exponential distribution with gamma priors when $a = p$ and $c = q$.

In the tables below, we present the results of numerical comparisons of the best fixed and fully sequential procedures for several values of the parameters. The tables depict the efficiency for the balanced and unbalanced designs. The results are presented for the comparisons of Poisson means with gamma priors.

Table 1: *Relative efficiency when $a = p$ and $c = q$.*

a	2.0001	2.0010	2.0100	2.1000	10	50	100	200
2.0001	0.0100	0.0152	0.0181	0.0189	0.0155	0.0150	0.0150	0.0149
2.0010	0.0152	0.0316	0.0479	0.0563	0.0483	0.0470	0.0468	0.0468
2.0100	0.0181	0.0479	0.0995	0.1481	0.1461	0.1431	0.1428	0.1426
2.1000	0.0189	0.0562	0.1481	0.3015	0.4021	0.3979	0.3973	0.3971
10	0.0155	0.0483	0.1461	0.4021	0.9428	0.9647	0.9670	0.9680
50	0.0150	0.0470	0.1431	0.3979	0.9647	0.9900	0.9923	0.9936
100	0.0150	0.0468	0.1428	0.3973	0.9670	0.9923	0.9949	0.9962
200	0.0149	0.0468	0.1426	0.3971	0.9680	0.9936	0.9962	0.9975

Table 2: *Relative efficiency when $\mu_1 = \mu_2$ and $\sigma_1 < \sigma_2$, $a = 2p$, $c = 2q$, and $a > c$.*

(a, c)	*Ratio*
(2.0001, 2.000001)	0.0048
(2.001, 2.0005)	0.0262
(2.01, 2.005)	0.0825
(2.1, 2.05)	0.2527
(10, 5)	0.9005
(50, 10)	0.9647
(100, 50)	0.9923
(200, 100)	0.9962

Table 3 gives the efficiency $\frac{r(\Delta_{\mathcal{S}})}{r(\Delta_{\mathcal{F}})}$ for the Poisson distribution with gamma priors when $a = p$ and $c = q$.

Table 3: *Relative efficiency when $a = p$ and $c = q$.*

a	10^{-10}	.001	.010	.100	1	10	50	100
10^{-10}	$2 * 10^{-5}$	.0280	.0874	.2475	.4431	.4938	.4988	.4994
0.001	.0280	.0560	.1154	.2755	.4711	.5218	.5267	.5274
0.010	.0874	.1154	.1748	.3349	.5305	.5812	.5868	.5868
0.100	.2475	.2755	.3349	.4950	.6906	.7413	.7463	.7469
1	.4431	.4711	.5305	.6906	.8862	.9369	.9419	.9425
10	.4938	.5218	.5812	.7413	.9369	.9876	.9925	.9932
50	.4988	.5267	.5868	.7463	.9419	.9925	.9975	.9981
100	.4994	.5274	.5868	.7469	.9425	.9933	.9981	.9988

The comparisons in Table 3 are for the balanced design.
Table 4 is given for $p = 10^{-10}$ with $a/p = c/q$ and $p > q$.
Table 6 gives the numerical values of the efficiency for $c = 4a$ and $q = 2p$.

Table 4: *Relative Efficiency when* $\mu_1 = \mu_2$ *and* $\sigma_1 < \sigma_2$.

(a, c)	*Ratio*
$(10^{-10}, 10^{-10})$	0.0000
$(0.0010, 10^{-5})$	0.0308
$(0.0100, 10^{-6})$	0.1333
$(0.1000, 10^{-7})$	0.4510
$(1.0000, 10^{-8})$	0.8591
$(10.0000, 10^{-9})$	0.9778
$(50.0000, 10^{-10})$	0.9931
$(100.0000, 10^{-10})$	0.9956

Table 5: *Relative Efficiency when* $\mu_1 < \mu_2$ *and* $\sigma_1 = \sigma_2$, $a = 2p$, $c = 2q$, *and* $a > c$.

(a, c)	*Ratio*
$(10^{-10}, 4 * 10^{-10})$	0.0000
(0.0010, 0.0040)	0.0885
(0.0100, 0.0400)	0.2694
(0.1000, 0.0400)	0.6513
(1.0000, 4.0000)	0.9349
(10.0000, 40.0000)	0.9930
(50.0000, 200.0000)	0.9903
(100.0000, 400.0000)	0.9993

Table 6: *Relative Efficiency when* $\mu_1 < \mu_2$ *and* $\sigma_1 < \sigma_2$, $a < c$, *and* $p < q$.

(a, c)	*Ratio*
$(10^{-10}, 4 * 10^{-10})$	0.0000
(0.0010, 0.0040)	0.1039
(0.0100, 0.0400)	0.3142
(0.1000, 0.0400)	0.6793
(1.0000, 4.0000)	0.9436
(10.0000, 40.0000)	0.9940
(50.0000, 200.0000)	0.9989
(100.0000, 400.0000)	0.9991

5 Concluding remarks

We have shown that the sequential procedure for the problem of estimating the difference of the means of two independent populations from the exponential family with conjugate priors when compared with the best fixed design reveal the superiority of the random design. The lower bound for the Bayes risk plus the expected costs determined. Application of the results to the Poisson and exponential distributions using gamma priors as well as the Bernoulli distribution with beta priors are given. Numerical comparisons of the best fixed and fully sequential procedures for several values of the parameters conducted. There are other random designs that are of interest including the two stage design, and the myopic design (see Terbeche, 2000). These designs seem to perform better than the best fixed design.

Acknowledgement

The authors wish to thank the referees for constructive criticisms which lead to substantial improvement of the paper.

References

Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd edition, New York: Springer-Verlag.

Diaconis, P. and Ylvisaker, D. (1979). Conjugate Priors for Exponential Families. *Ann. Statist.*, 6, 269-281.

Hajek, J. & Sidak. Z. (1967). *Theory of Rank Tests*, New York-London: Academic Press.

Lehmann, E.L. (1959). *Testing Statistical Hypothesis*, New York: John Wiley and Sons Inc.

Terbeche, M. (2000). *Sequential Design for Estimation*, Thesis, Florida Institute of Technology.

West, M. (1985). *Generalized Linear Models: Outlier Accommodation, Scale Parameters and Prior Distributions*. Bayesian Statistics 2, J.M. Bernardo, M.H. De Groot, D.V. Lindley and A.F.M. Smith (Eds), North-Holland, Amsterdam and Valencia University Press.

West, M. (1986). Bayesian Model Monitoring. *J. Roy. Statist. Soc.*, (Ser. B) 28, 70-78.